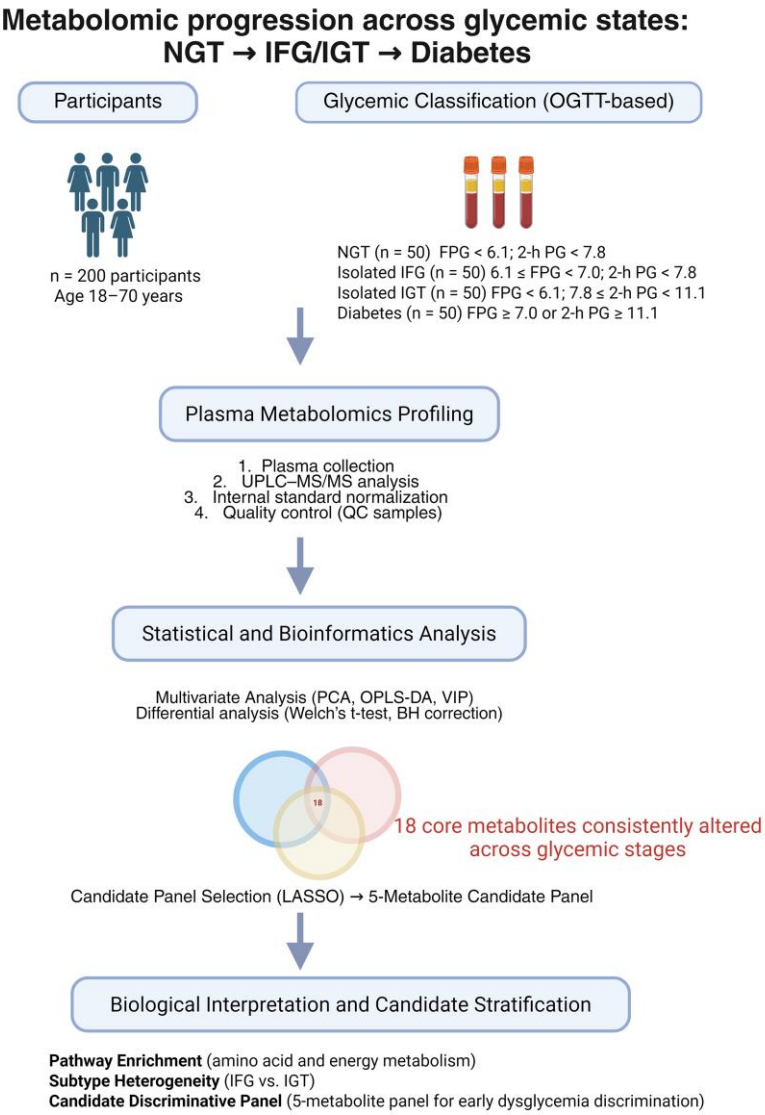
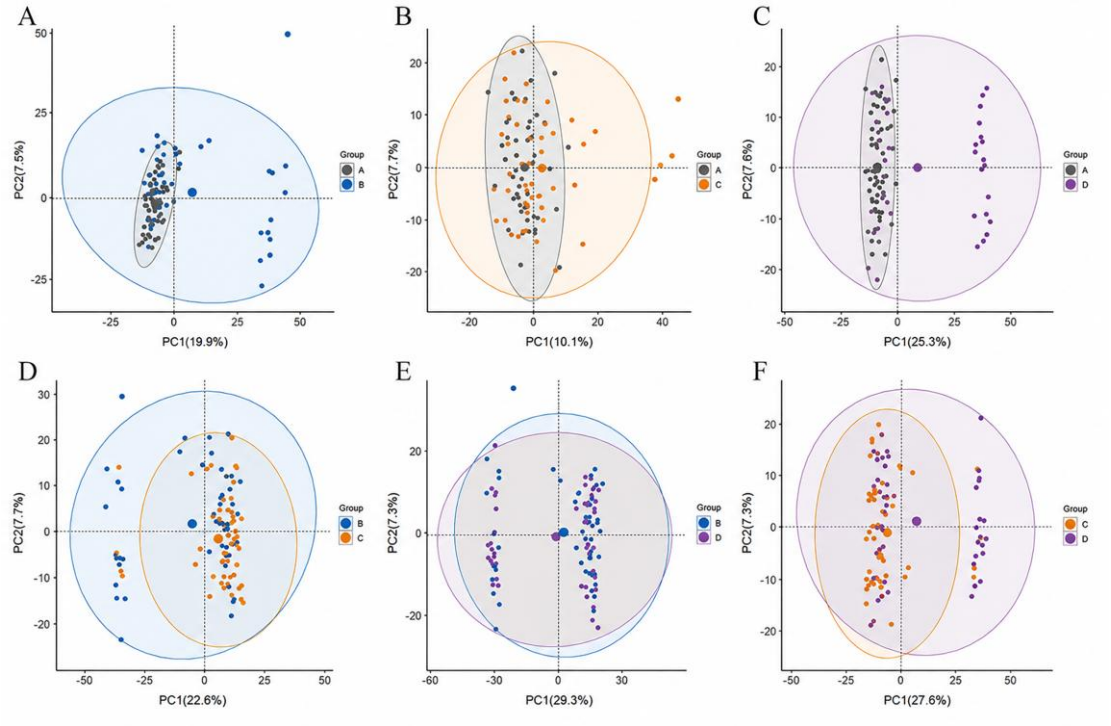
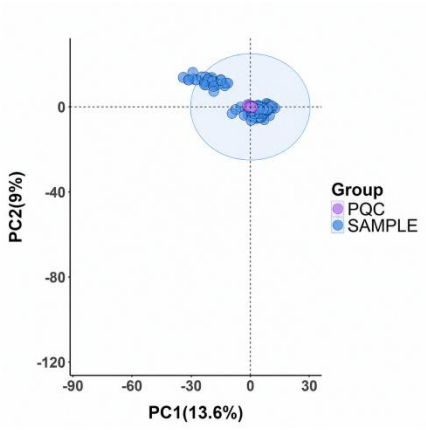


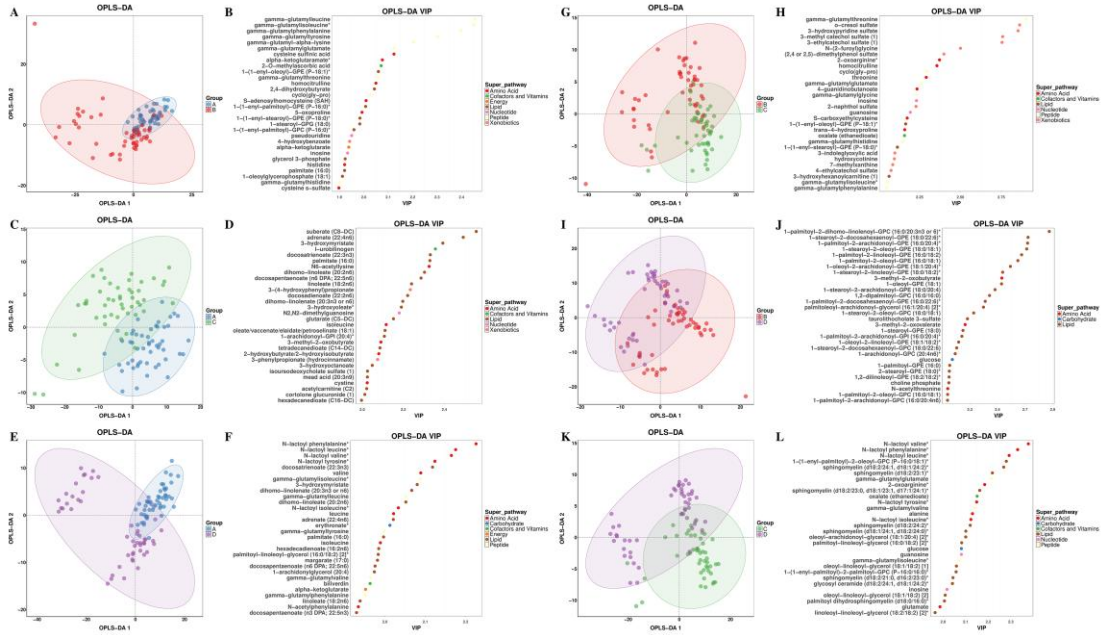
Supplementary Figure 1 Study design and analytical workflow of the present study. The schematic summarizes the study design and analytical workflow across the glycemic continuum from normal glucose tolerance to diabetes, including oral glucose tolerance test-based grouping, plasma metabolomics profiling, pathway enrichment, subtype heterogeneity analysis, and candidate discriminative panel selection. NGT: Normal glucose tolerance; IFG: Impaired fasting glucose; IGT: Impaired glucose tolerance; OGTT: Oral glucose tolerance test; UPLC-MS/MS: Ultra-performance liquid chromatography-tandem mass spectrometry; LASSO: Least absolute shrinkage and selection operator.



Supplementary Figure 2 Principal component analysis score plots of pairwise comparisons among study groups. A: Principal component analysis (PCA) score plot for the comparison between the normal glucose tolerance (NGT) and isolated impaired fasting glucose (IFG) groups; B: PCA score plot for the comparison between the NGT and isolated impaired glucose tolerance (IGT) groups; C: PCA score plot for the comparison between the NGT and diabetes groups; D: PCA score plot for the comparison between the isolated IFG and isolated IGT groups; E: PCA score plot for the comparison between the isolated IFG and diabetes groups; F: PCA score plot for the comparison between the isolated IGT and diabetes groups. Each point represents an individual sample, and ellipses indicate the confidence regions for each group. The separation and clustering patterns in the PCA space reflect differences in metabolic profiles among groups.

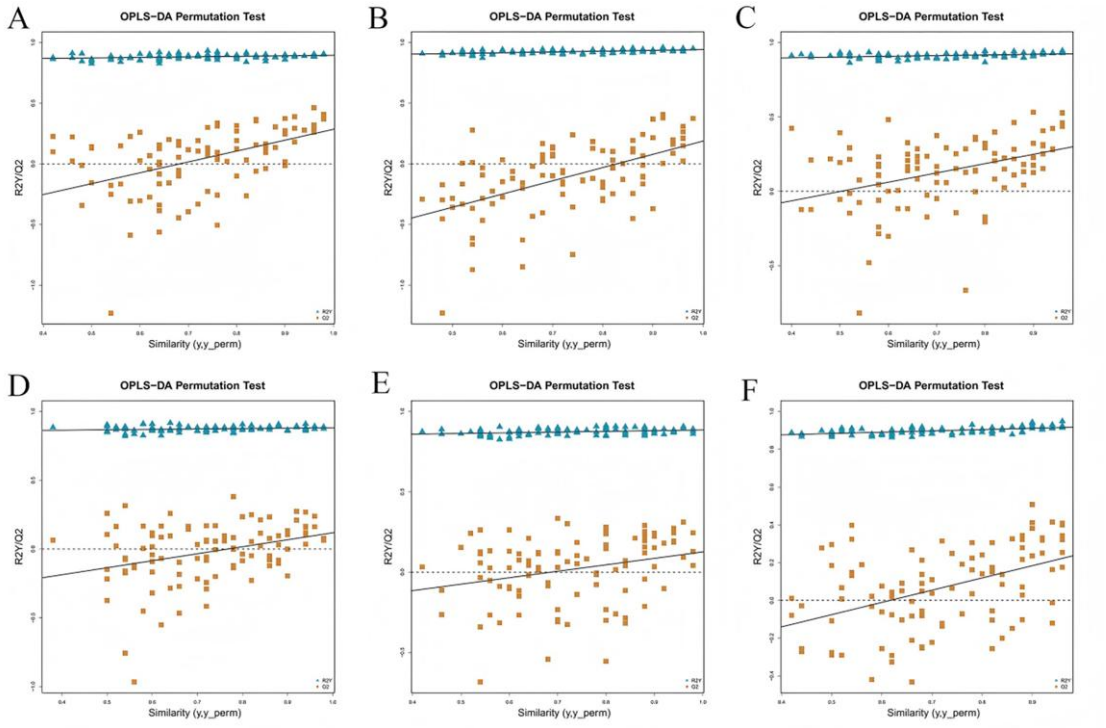


Supplementary Figure 3 Principal component analysis score plot for quality control assessment. The principal component analysis score plot shows the distribution of pooled quality control samples and study samples based on global metabolomic profiles. Pooled quality control samples (pooled quality control samples, purple dots), generated by mixing equal volumes of all study samples, were tightly clustered, indicating good analytical reproducibility and stability of the ultra-performance liquid chromatography-tandem mass spectrometry platform. Study samples (blue dots) showed broader dispersion, reflecting inter-individual biological variability. One potential outlier was observed but did not materially affect the overall clustering pattern. PQC: Pooled quality control.



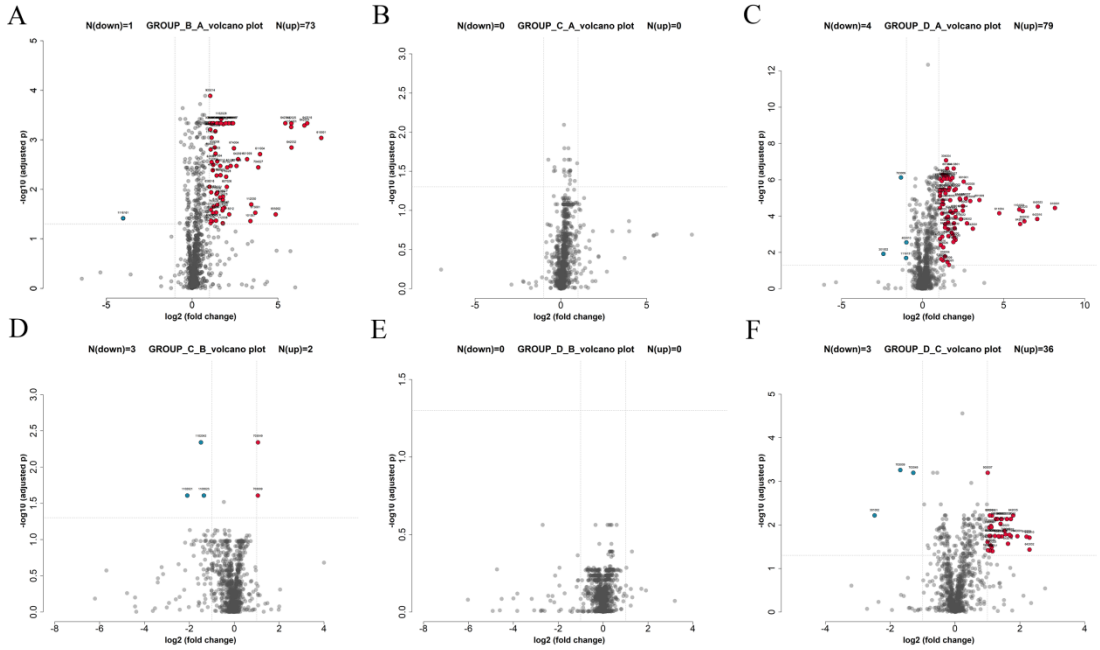
Supplementary Figure 4 Orthogonal partial least squares-discriminant analysis score plots and variable importance in projection analyses for pairwise comparisons among glycemic groups. A and B: Orthogonal partial least squares-discriminant analysis (OPLS-DA) score plot and variable importance in projection (VIP) plot for the comparison between the normal glucose tolerance (NGT) group (group A) and the isolated impaired fasting glucose (IFG) group (group B); C and D: OPLS-DA score plot and VIP plot for the comparison between the NGT group (group A) and the isolated IGT group (group C); E and F: OPLS-DA score plot and VIP plot for the comparison between the NGT group (group A) and the diabetes group (group D); G and H: OPLS-DA score plot and VIP plot for the comparison between the isolated IFG group (group B) and the isolated IGT group (group C); I and J: OPLS-DA score plot and VIP plot for the comparison between the isolated IFG group (group B) and the diabetes group (group D); K and L: OPLS-DA score plot and VIP plot for the comparison between the isolated IGT group (group C) and the diabetes group (group D). In the OPLS-DA score plots, each point represents an individual sample, and ellipses indicate the confidence regions for each group. In the VIP plots, metabolites are ranked according to their VIP scores, and different colors represent different

metabolic categories.

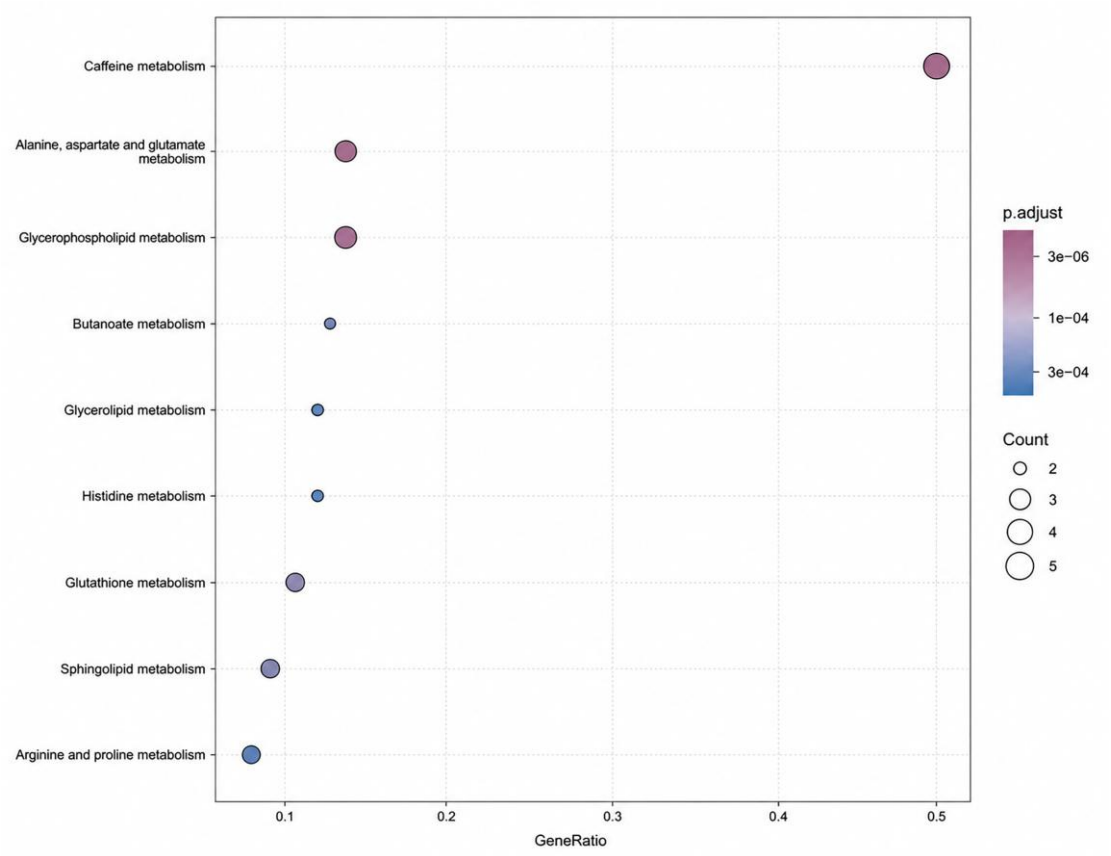


Supplementary Figure 5 Permutation tests of the orthogonal partial least squares-discriminant analysis models for pairwise comparisons among study groups. A: Group A *vs* group B [normal glucose tolerance (NGT) *vs* isolated impaired fasting glucose (IFG)]; B: Group A *vs* group C [NGT *vs* isolated impaired glucose tolerance (IGT)]; C: Group A *vs* group D (NGT *vs* diabetes); D: Group B *vs* group C (isolated IFG *vs* isolated IGT); E: Group B *vs* group D (isolated IFG *vs* diabetes); F: Group C *vs* group D (isolated IGT *vs* diabetes). Permutation tests were performed to evaluate the robustness of the orthogonal partial least squares-discriminant analysis models and the risk of overfitting. The X-axis represents the similarity between the original and permuted class labels, and the y-axis shows the corresponding explained variation in the response variable and predictive ability of the model estimated by cross-validation (Q2) values. The regression lines of explained variation in the response variable and Q2 are displayed. A negative intercept of the Q2 regression line indicated that the models were not obviously overfitted. OPLS-DA: Orthogonal partial least squares-discriminant analysis;

R2Y: Explained variation in the response variable; Q2: Predictive ability of the model estimated by cross-validation.



Supplementary Figure 6 Volcano plots of differential metabolites among different glycemic states. A: Volcano plot for the comparison between the normal glucose tolerance (NGT) and isolated impaired fasting glucose (IFG) groups; B: Volcano plot for the comparison between the NGT and isolated impaired glucose tolerance (IGT) groups; C: Volcano plot for the comparison between the NGT and diabetes groups; D: Volcano plot for the comparison between the isolated IFG and isolated IGT groups; E: Volcano plot for the comparison between the isolated IFG and diabetes groups; F: Volcano plot for the comparison between the isolated IGT and diabetes groups. Each point represents a metabolite. The X-axis indicates $\log_2$ fold change, and the Y-axis indicates -$\log_{10}$(adjusted *P* value). Colored points represent significantly altered metabolites, and gray points represent non-significant metabolites. The numbers of upregulated and downregulated metabolites are indicated in each panel.



Supplementary Figure 7 Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis of differential metabolites between the isolated impaired fasting glucose and isolated impaired glucose tolerance groups. Bubble plot showing the Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis of differential metabolites between the isolated impaired fasting glucose and isolated impaired glucose tolerance groups. The X-axis represents the enrichment ratio, and the Y-axis indicates the enriched metabolic pathways. Bubble size corresponds to the number of differential metabolites enriched in each pathway, and bubble color indicates the significance level. Pathways with a higher enrichment ratio and lower *P* value were considered more significantly enriched.

Supplementary Table 1 Validation metrics of pairwise OPLS-DA models across glycemic states

Comp	Groups	n/n	R^2X(cu	R^2Y(cu	Q^2(cu	CV-ANO	CV-ANOV	CV-ANOVA *P*

arison	compared		m)	m)	m)	VA F	A df	value
B vs A	IFG vs NGT	50/50	0.158	0.330	0.270	17.94	(2, 97)	2.35e-07
C vs A	IGT vs NGT	50/50	0.187	0.827	0.244	5.00	(6, 93)	1.75e-04
D vs A	Diabetes vs NGT	50/50	0.296	0.801	0.458	13.10	(6, 93)	1.09e-10
C vs B	IGT vs IFG	50/50	0.152	0.205	0.081	4.27	(2, 97)	1.67e-02
D vs B	Diabetes vs IFG	50/50	0.290	0.672	-0.016	N/A	N/A	N/A
D vs C	Diabetes vs IGT	50/50	0.258	0.793	0.231	4.66	(6, 93)	3.51e-04

Pairwise OPLS-DA models were constructed for all pairwise comparisons among the four glycemic groups. The n/n column indicates the sample size in each compared group. $R^2X(cum)$, $R^2Y(cum)$, and $Q^2(cum)$ represent cumulative explained variance in X, cumulative explained variance in Y, and cumulative predictive ability, respectively. Model significance was assessed using CV-ANOVA. N/A indicates that the CV-ANOVA result was not available for the corresponding model. Abbreviations: CV-ANOVA, cross-validated analysis of variance; IFG, impaired fasting glucose; IGT, impaired glucose tolerance; NGT, normal glucose tolerance; OPLS-DA, orthogonal partial least-squares discriminant analysis.

Supplementary Table 2 Differential abundance and confounder-adjusted trend statistics for the 18 shared core metabolites across glycemic states

Meta bolite	Main path way	Sub pathway	VIP rang e	5-me tabol ite	B/A FC (FD	C/A FC (FDR)	D/A FC (FDR)	Adjuste d trend β (95%	*P* for trend	FD R for

				panel	**R)**			**CI)**		**trend**
isoleucine	Amino Acid	Leucine, Isoleucine and Valine Metabolism	1.68 – 2.56	No	1.131 (0.003)	1.156 (0.016)	1.180 (2.20e-05)	0.156 (0.040 to 0.271)	0.008	0.011
leucine	Amino Acid	Leucine, Isoleucine and Valine Metabolism	1.90 – 2.47	Yes	1.212 (4.62e-04)	1.153 (0.022)	1.245 (1.48e-06)	0.198 (0.076 to 0.321)	0.002	0.003
valine	Amino Acid	Leucine, Isoleucine and Valine Metabolism	1.94 – 2.36	Yes	1.135 (5.77e-04)	1.139 (0.03)	1.192 (1.21e-06)	0.183 (0.062 to 0.304)	0.003	0.005
N6-acetyllysine	Amino Acid	Lysine Metabolism	1.49 – 3.02	No	1.269 (0.01)	1.138 (0.008)	1.132 (2.67e-04)	0.168 (0.026 to 0.310)	0.02	0.022
spermine	Amino Acid	Polyamine Metabolism	1.13 – 2.37	No	1.454 (0.022)	1.606 (0.013)	1.444 (0.005)	0.186 (0.044 to 0.328)	0.011	0.013
2-O-	Cofact	Ascorbat	1.89	No	1.856	1.466	2.113	0.147	0.03	0.03

methylascorbic acid	ors and Vitamins	e and Aldarate Metabolism	– 2.17		(5.99e-04)	(0.041)	(3.64e-06)	(0.015 to 0.280)		
alpha-ketoglutarate	Energy	TCA Cycle	2.04 – 2.11	No	2.088 (6.21e-04)	1.44 (0.032)	2.694 (5.73e-07)	0.243 (0.121 to 0.365)	1.20e-04	7.17e-04
glutarate (C5-DC)	Lipid	Fatty Acid, Dicarboxylate	1.96 – 2.36	Yes	1.937 (0.025)	1.897 (4.62e-04)	2.830 (8.84e-07)	0.249 (0.114 to 0.384)	3.67e-04	0.002
2-hydroxypalmitate	Lipid	Fatty Acid, Monohydroxy	1.93 – 2.00	No	2.393 (4.62e-04)	1.440 (0.046)	2.547 (3.49e-06)	0.205 (0.069 to 0.342)	0.003	0.005
3-hydroxymyristate	Lipid	Fatty Acid, Monohydroxy	2.07 – 2.09	No	1.334 (0.02)	1.431 (0.031)	1.743 (2.26e-07)	0.307 (0.178 to 0.435)	5.28e-06	9.51e-05
adrenate (22:4n6)	Lipid	Long Chain Polyunsaturated Fatty Acid (n3 and n6)	1.97 – 2.00	No	1.553 (0.003)	1.427 (0.036)	1.858 (5.66e-07)	0.238 (0.101 to 0.374)	7.28e-04	0.003
arachidonat	Lipid	Long Chain	1.92 –	No	3.511 (4.80	1.756 (0.042)	4.122 (3.26e-	0.220 (0.085	0.002	0.003

e (20:4n6)		Polyunsaturated Fatty Acid (n3 and n6)	2.07		e-04)		06)	to 0.356)		
docosapentaenoate (n6 DPA; 22:5n6)	Lipid	Long Chain Polyunsaturated Fatty Acid (n3 and n6)	2.06 – 2.07	No	1.797 (0.001)	1.476 (0.031)	2.362 (5.22e-07)	0.279 (0.145 to 0.412)	6.11e-05	5.50e-04
palmitate (16:0)	Lipid	Long Chain Saturated Fatty Acid	1.93 – 1.99	Yes	1.355 (3.81e-04)	1.192 (0.036)	1.476 (7.81e-07)	0.217 (0.084 to 0.349)	0.002	0.003
adenine	Nucleotide	Purine Metabolism, Adenine containing	1.56 – 1.98	No	2.388 (0.022)	1.341 (0.032)	2.876 (8.77e-05)	0.222 (0.084 to 0.360)	0.002	0.003
phenylalanylglycine	Peptide	Dipeptide	1.49 – 1.96	No	0.591 (0.004)	0.795 (0.031)	0.407 (2.67e-04)	-0.166 (-0.300 to -0.031)	0.016	0.018
threonylph	Peptide	Dipeptide	1.51 –	Yes	0.402 (4.62	0.676 (0.028)	0.452 (3.23e-	-0.192 (-0.327	0.005	0.008

enylalanine			2.20 – 2.27		e-04)		04)	to -0.057)		
glu-gly-asn-val	Peptide	Polypeptide	1.69 – 2.27	No	0.754 (0.032)	0.917 (0.002)	0.342 (7.01e-05)	-0.227 (-0.361 to -0.092)	0.001	0.003

B/A, C/A, and D/A FCs represent mean abundance FCs for isolated IFG, isolated IGT, and diabetes compared with NGT, respectively. VIP range represents the range of variable importance in projection values across the three main OPLS-DA comparisons: A vs B, A vs C, and A vs D. FDR values were calculated using the Benjamini-Hochberg method. Adjusted trend β values were estimated using multivariable linear regression with standardized log2-transformed metabolite abundance as the dependent variable and glycemic status coded ordinally from NGT to diabetes. Models were adjusted for age, sex, BMI, TG, SBP, DBP, ALT, and AST. The 5-metabolite panel column indicates whether the metabolite was selected by LASSO regression. Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; CI, confidence interval; DBP, diastolic blood pressure; FC, fold change; FDR, false discovery rate; IFG, impaired fasting glucose; IGT, impaired glucose tolerance; LASSO, least absolute shrinkage and selection operator; NGT, normal glucose tolerance; OPLS-DA, orthogonal partial least squares-discriminant analysis; SBP, systolic blood pressure; TCA, tricarboxylic acid; TG, triglycerides; VIP, variable importance in projection.

Supplementary Table 3 Incremental discriminative performance of the five-metabolite candidate panel over standard clinical indicators for three pairwise comparisons

Comparison	**AUC (Base**	**AUC**	**AUC *P***	**NRI**	**NRI *P***	**IDI**	**IDI**	***P***

	model, 95%CI)	(Augmented model, 95% CI)	value	(95% CI)	value	(95% CI)	value
A *vs* B	0.663 (0.547, 0.779)	0.868 (0.791, 0.944)	< 0.001	1.032 (0.649, 1.389)	< 0.001	0.323 (0.226, 0.420)	< 0.001
A *vs* C	0.755 (0.657, 0.853)	0.868 (0.795, 0.941)	0.011	0.751 (0.353, 1.133)	< 0.001	0.207 (0.123, 0.291)	< 0.001
A *vs* D	0.842 (0.752, 0.932)	0.945 (0.899, 0.990)	0.019	1.420 (1.120, 1.711)	< 0.001	0.283 (0.179, 0.386)	< 0.001

Group A: Normal glucose tolerance; Group B: Isolated impaired fasting glucose; Group C: Isolated impaired glucose tolerance; Group D: Diabetes. Data are presented for three pairwise comparisons: (1) A *vs* B; (2) A *vs* C; and (3) A *vs* D. The base model included standard clinical indicators and potential confounders, including age, sex, body mass index, triglycerides, systolic blood pressure, diastolic blood pressure, alanine aminotransferase, and aspartate aminotransferase. The augmented model additionally incorporated the five selected metabolites. Discriminative performance was evaluated using area under the curve with 95%CI. The incremental discriminative value of the five-metabolite candidate panel was assessed using continuous net reclassification improvement and integrated discrimination improvement, with corresponding 95%CIs and *P* values. AUC: Area under the curve; NRI: Net reclassification improvement; IDI: Integrated discrimination improvement.

Supplementary Table 4 Variance inflation factor analysis of the final five-metabolite logistic regression model

Variable	VIF
Leucine	4.266
Valine	4.186
Glutarate (C5-DC)	1.086
Palmitate (16:0)	1.304
Threonylphenylalanine	1.205

VIFs were calculated to assess multicollinearity among predictors in the final five-metabolite logistic regression model. All variance inflation factor values were below 5, indicating no substantial multicollinearity among the selected metabolites. VIF: Variance inflation factor; C5-DC: Glutarate.

Supplementary Table 5 Associations between weighted gene co-expression network analysis-derived metabolite modules and isolated impaired fasting glucose/impaired glucose tolerance status

Module	Size	Correlation	*P* value	FDR
MEturquoise	167	-0.279	0.005	0.024
MEpink	35	-0.268	0.007	0.024
MEbrown	117	0.264	0.008	0.024
MEblue	144	0.151	0.133	0.247
MEyellow	113	-0.15	0.137	0.247
MEgreen	92	-0.062	0.543	0.733
MEblack	44	-0.050	0.624	0.733
MEgrey	217	-0.046	0.652	0.733
MEred	62	0.005	0.964	0.964

Module eigengenes were correlated with impaired fasting glucose (IFG)/impaired glucose tolerance (IGT) status coded as isolated IFG = 0 and isolated IGT = 1. Positive and negative correlations indicate higher module eigengene values in isolated IGT and isolated IFG, respectively. *P* values were adjusted using the Benjamini-Hochberg method, and false discovery rate <

0.05 was considered statistically significant. FDR: False discovery rate.

Supplementary Table 6 Functional annotation of significant weighted gene co-expression network analysis modules associated with isolated impaired fasting glucose/impaired glucose tolerance status

Module	**Direction**	**Size**	**Correlation**	***P* value**	**FDR**	**Major super meta pathways**	**Top sub meta pathways**
turquoise	Higher in isolated IFG	167	-0.279	0.005	0.024	Lipid (73); Amino Acid (33); Peptide (16); Carbohydrate (11); Nucleotide (10)	Diacylglycerol (14); Gamma-glutamyl Amino Acid (12); Lysophospholipid (12); Fatty Acid Metabolism (Acyl Choline) (9); Partially Characterized Molecules (8); Glutamate Metabolism (7); Fatty Acid, Dicarboxylate (5); Fatty Acid, Monohydroxy (5)
pink	Higher in isolated IFG	35	-0.268	0.007	0.024	Xenobiotics (31); Amino Acid (3); Energy (1)	Xanthine Metabolism (12); Food Component/Plant (8); Benzoate Metabolism (5); Chemical (4); Tobacco Metabolite (2); Histidine Metabolism (1); Methionine, Cysteine, SAM and Taurine Metabolism (1); TCA Cycle (1)

brown	Higher in isolated IGT	117	0.264	0.008	0.024	Lipid (45); Amino Acid (21); Peptide (20); Cofactors and Vitamins (12); Nucleotide (12)	Fibrinogen Cleavage Peptide (12); Fatty Acid Metabolism (Acyl Carnitine, Hydroxy) (7); Fatty Acid Metabolism (Acyl Carnitine, Monounsaturated) (7); Phosphatidylcholine (PC) (6); Fatty Acid Metabolism (Acyl Carnitine, Medium Chain) (5); Ascorbate and Aldarate Metabolism (4); Dipeptide (4); Glutathione Metabolism (4)

Module eigengenes were correlated with group status coded as isolated impaired fasting glucose = 0 and isolated impaired glucose tolerance = 1. Positive correlations indicate higher module eigengene values in isolated impaired glucose tolerance, whereas negative correlations indicate higher module eigengene values in isolated impaired fasting glucose. Significant modules were defined as those with false discovery rate < 0.05. Major super meta pathways and top sub meta pathways were summarized according to metabolite annotations within each significant module, with numbers in parentheses indicating the number of metabolites assigned to each pathway category. IFG: Impaired fasting glucose; IGT: Impaired glucose tolerance; FDR: False discovery rate; TCA: Tricarboxylic acid; PC: Phosphatidylcholine; SAM: S-adenosylmethionine.